Available online at www.sciencedirect.com
SCIENCE DIRECT
Discrete Applied Mathematics 154 (2006) 493–507
DISCRETE APPLIED MATHEMATICS
www.elsevier.com/locate/dam

Description-meet compatible multiway dissimilarities

Jean Diatta*

IREMIA, Université de la Réunion, 15 avenue René Cassin - BP 7151, 97715 Saint-Denis messag cedex 9, France

Received 19 September 2002; received in revised form 8 March 2005; accepted 20 July 2005
Available online 3 October 2005

Abstract

Multiway dissimilarities are a natural generalization of standard pairwise ones, that allow global comparison of more than two entities. Assuming the entity descriptions belong to a complete meet-semilattice, we consider so-called description-meet compatible multiway dissimilarities on the entity set; that is, multiway dissimilarities agreeing with entity descriptions in the following sense: the lower the greatest lower bound of the descriptions of entities in a given subset, the more dissimilar the entities in this subset. On the one hand, we show that when the entity description set is of breadth k, strictly description-meet compatible k-way dissimilarities are quasi-ultrametric. By duality, when entity descriptions belong to a complete join-semilattice, a similar result holds for so-called strictly description-join compatible multiway dissimilarities. Moreover, we study relationships between multiway dissimilarities in general, and provide examples of description-meet compatible ones.

Keywords: Description-meet compatibility; Multiway dissimilarity; Quasi-ultrametric

1. Introduction

Measures of comparison play an important role in many domains including pattern recognition, case based reasoning, clustering, classification, and machine learning. Dissimilarities (or, dually, similarities) are among the most studied and most used measures of comparison. However, they suffer from one practical limitation. Indeed, they allow only pairwise comparison of entities, although many applications often require capturing global (dis)similarity degree of more than two entities. Now, it is a fact that the actual (dis)similarity degree of an entity set is seldom expressible in terms of pairwise (dis)similarity degrees of entities in the set. It is then interesting to have tools for globally assigning a (dis)similarity degree to any entity subset. Multiway dissimilarities are such tools. They are a natural generalization of classical (two-way) dissimilarities. In the last decade, they have been investigated or considered from different approaches by many authors among whom we just mention Batbedat [3], Bandelt and Dress [2], Joly and Le Calvé [15], Daws [8], Bennani and Heiser [4], Diatta [10].

As (dis)similarity degrees are necessarily computed on the basis of entity descriptions, it is hopeful that the measure of comparison resulting from these computations agree with these descriptions as well as possible. Clearly, the manner of taking into account entity descriptions will depend on the purposes intended to the measure of comparison. In this paper, we are concerned with multiway dissimilarities on entities whose descriptions are assumed to belong to a complete meet-semilattice. These entity descriptions are then taken into account through a condition called description-meet

* Tel.: +33 262 938303; fax: +33 262 938360.
E-mail address: jean.diatta@univ-reunion.fr.

0166-218X/$ - see front matter
doi:10.1016/j.dam.2005.07.013

compatibility; that is, a kind of natural agreement between the (multiway) dissimilarity and the entity descriptions, expressing the following fact: the lower the greatest lower bound of the descriptions of entities in a given subset, the more dissimilar entities in this subset. Multiway dissimilarities satisfying this agreement condition are said to be description-meet compatible.

We show that when the entity description set is of breadth k, strictly description-meet compatible k-way dissimilarities are quasi-ultrametric. By duality, when the entity descriptions belong to a complete join-semilattice, a similar result holds for so-called strictly description-join compatible multiway dissimilarities. Now, quasi-ultrametric k-way dissimilarities are known to be in bijection with the so-called indexed k-quasi-hierarchies [10] also known as indexed closed weak hierarchies of breadth at most k [2] or k-weak hierarchical representations [6]. On the other hand, we study relationships between multiway dissimilarities in general, and provide examples of description-meet compatible ones, among which, a generalization of the well-known Ochiaï's dissimilarity. The paper is organized as follows.

Section 2 introduces multiway dissimilarities after recalling classical pairwise ones. It is also shown there that a $2k$-point implication introduced in [2] is equivalent to the conjunction of a diameter and inclusion conditions characterizing so-called quasi-ultrametric k-way dissimilarities. Description-meet compatibility is addressed in Section 3 and relationships between multiway dissimilarities are studied in Section 4. Finally, examples of description-meet compatible multiway dissimilarities are provided in Section 5, and a short conclusion closes the paper.

2. Multiway dissimilarities

2.1. From pairwise to multiway dissimilarities

Before introducing multiway dissimilarities, let us first recall the classical pairwise ones. Let E be a finite nonempty set.

A (pairwise) dissimilarity on E is a map $d : E \times E \to \mathbb{R}$ satisfying reflexivity ((R2) $d(x, x) = 0$), non-negativity ((N2) $d(x, y) \geqslant 0$) and symmetry ((S2) $d(x, y) = d(y, x)$).

Considering maps on $E^3, E^4, \ldots, E^k$, with similar properties, naturally leads to the notion of 3-way, 4-way, ..., k-way dissimilarity. For instance, a 3-way dissimilarity on E will be any map $d : E^3 \to \mathbb{R}$ satisfying: (R3) $d(x, x, x) = 0$, (N3) $d(x, y, z) \geqslant 0$ and (S3) $d(x, y, z) = d(x, z, y) = d(y, x, z) = d(y, z, x) = d(z, x, y) = d(z, y, x)$. The term *multiway* dissimilarity will be used to mean a k-way dissimilarity, for some $k \geqslant 2$.

Of course, due to the tuple-based definition above, the complexity of expressions related to k-way dissimilarities increases when k grows. Then, for the sake of simplicity, we adopt in the present paper a set-based definition based on the following observation: according to (R2) and (N2), $d(x, x) \leqslant d(x, y)$ for all x, y. Thus, a dissimilarity on E can be defined as being a nonnegative real valued map d on the set of singletons and pairs of E, satisfying $d(\{x\}) = 0$ and $d(\{x\}) \leqslant d(\{x, y\})$. This set-based definition makes the symmetry condition implicit. Moreover, for $k \geqslant 2$, its generalization to k-way dissimilarities involves shortest expressions.

For reasons explained in Remark 13 below, we will drop out the reflexivity condition and thus be rather concerned with so-called (multiway) pseudo-dissimilarities. However, we will still use the term dissimilarity, keeping in mind that the condition $d(\{x\}) = 0$ is not required.

For any set S and any integer $k \geqslant 1$, $S^*_{\leqslant k}$ will denote the set of all nonempty subsets of S with at most k elements. Then, we formally define multiway dissimilarities as follows.

Definition 1. A k-way dissimilarity on E will be any nonnegative real valued and isotone map defined on the set of all nonempty subsets of E with at most k elements, i.e., any map $d : E^*_{\leqslant k} \to \mathbb{R}_+$ such that $d(X) \leqslant d(Y)$ when $X \subseteq Y$.

Example 2. Table 1 presents a dataset about seven market baskets and five items: bread (brd), butter (btr), cheese (chs), eggs (egg), milk (mlk); for instance, the market basket labeled 1 contains bread and cheese. For any k such that $2 \leqslant k \leqslant 5$, a k-way dissimilarity on the set of items, can be defined by letting $\mathrm{dis}_k(X)$ be seven minus the number of baskets that contain each of the items in X. Then, for instance, $\mathrm{dis}_3(\{\mathrm{brd}, \mathrm{chs}\}) = 4$ and $\mathrm{dis}_3(\{\mathrm{brd}, \mathrm{btr}, \mathrm{chs}\}) = 7$.

Remark 3. For $\{x, y, z\} \subseteq E$, we will simply write $d(x)$ or $d(x, y)$ or $d(x, y, z)$ instead of $d(\{x\})$ or $d(\{x, y\})$ or $d(\{x, y, z\})$, respectively. Moreover, as in the tuple-based setting, the notation $d(x, y)$ or $d(x, y, z)$ will not require x, y and z be distinct.

Table 1
Example dataset

	brd	btr	chs	egg	mlk
1	x		x		
2		x	x		x
3	x		x	x	
4		x		x	x
5	x	x			x
6	x	x		x	
7	x		x		x

2.2. Max-extension and canonical restriction

From any (pairwise) dissimilarity can be derived various 3-way ones. Among these 3-way dissimilarities is the so-called 3-way max-extension: the 3-way *max-extension* of a (2-way) dissimilarity d_2 is the 3-way dissimilarity d_3 defined by $d_3(x,y,z)=\max\{d_2(x,y),d_2(x,z),d_2(y,z)\}$.

Example 4. Let $\mathrm{dis}_{2,3}$ denote the 3-way max-extension of the dissimilarity dis_2 defined in Example 2. Then

$$\begin{aligned}\mathrm{dis}_{2,3}(\mathrm{brd},\mathrm{btr},\mathrm{chs}) = {} & \max\{\mathrm{dis}_2(\mathrm{brd},\mathrm{btr}),\mathrm{dis}_2(\mathrm{brd},\mathrm{chs}),\\ & \mathrm{dis}_2(\mathrm{btr},\mathrm{chs})\}=\max\{5,4,6\}=6.\end{aligned}$$

This shows that dis_3 is not the 3-way max-extension of dis_2 since $\mathrm{dis}_3(\mathrm{brd},\mathrm{btr},\mathrm{chs})=7$.

More generally, for two integers k and l such that $2\leqslant k\leqslant l$, the l-way max-extension of a k-way dissimilarity d_k is the l-way dissimilarity d_l defined by $d_l(X)=\max_{Y\in X^*_{\leqslant k}} d_k(Y)$.

Conversely, from any 3-way dissimilarity can be derived various 2-way ones. Among these 2-way dissimilarities is the so-called 2-way canonical restriction: the 2-way *canonical restriction* of an 3-way dissimilarity d_3 is the 2-way dissimilarity d_2 defined by $d_2(x,y)=d_3(x,y)$.

Example 5. For two integers k and l such that $2\leqslant k\leqslant l\leqslant 5$, let $\mathrm{dis}_{l,k}$ denote the k-way canonical restriction of the l-way dissimilarity dis_l defined in Example 2. Then for any such k and l, $\mathrm{dis}_{l,k}=\mathrm{dis}_k$.

More generally, for two integers k and l such that $2\leqslant k\leqslant l$, the k-way canonical restriction of an l-way dissimilarity d_l is the k-way dissimilarity d_k defined by $d_k(X)=d_l(X)$.

2.3. Quasi-ultrametric multiway dissimilarities

Key notions in the definition of quasi-ultrametrics given below are those of a d-ball, (d,k)-ball and d-diameter, where d is a k-way dissimilarity. To catch their meaning, let us first cast them in the case of a pairwise dissimilarity, say d_2.

The d_2-diameter of a nonempty subset Z of E is the maximum d_2-dissimilarity degree between elements of Z, i.e., $\mathrm{diam}_{d_2}(Z)=\max\{d_2(x,y):x,y\in Z\}$.

Let now x and y be two distinct elements of E and let r be a nonnegative real number. The d_2-ball of center x and radius r is the set $B^{d_2}(x,r)$ of elements of E whose d_2-dissimilarity degree from $\{x\}$ is at most r, i.e., formally, $B^{d_2}(x,r)=\{z\in E: d_2(x,z)\leqslant r\}$; the $(d_2,2)$-ball (or 2-ball) generated by x is the set $B^{d_2}_x=B^{d_2}(x,d_2(x))$, and the $(d_2,2)$-ball generated by $\{x,y\}$ is the set $B^{d_2}_{xy}=B^{d_2}(x,d_2(x,y))\cap B^{d_2}(y,d_2(x,y))$. If $x=y$, $B^{d_2}_{xy}=B^{d_2}_x$.

Fig. 1 illustrates the notions of a ball and a 2-ball, in the case of an Euclidean dissimilarity.

All these notions have been naturally generalized to multiway dissimilarities in [10]. For $k\geqslant 2$, let d_k denote a k-way dissimilarity on E. For any subset S of E and any element $x\in E$, $S+x$ will denote $S\cup\{x\}$; similarly, $S-x$ will denote $S\backslash\{x\}$.

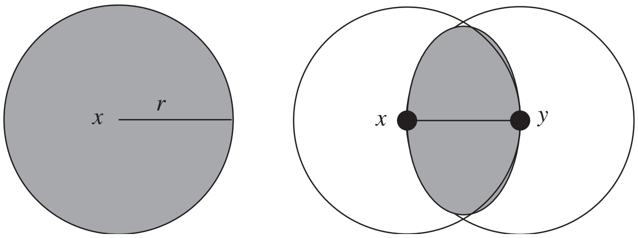

Fig. 1. d_2-ball of center x and radius r and $(d_2, 2)$-ball generated by $\{x, y\}$.

The d_k-*diameter* (or, simply, *diameter*) of a nonempty subset Z of E is the maximum d_k-dissimilarity degree between elements of Z, i.e., $\mathrm{diam}_{d_k}(Z) = \max\{d_k(T) : T \in Z^*_{\leqslant k}\}$.

Let $X \in E^*_{\leqslant k-1}$. The d_k-*ball* (or, simply, *ball*) of center X and radius r is the set $B^{d_k}(X, r)$ defined by $B^{d_k}(X, r) = \{y \in E : d_k(X + y) \leqslant r\}$. If $X \in E^*_{\leqslant k}$, then the (d_k, k)-*ball* (or, simply, k-*ball* relative to d_k) generated by X will be the set $B^{d_k}_X$ defined by $B^{d_k}_X = B^{d_k}(X, d_k(X))$ when $|X| \leqslant k-1$, and $B^{d_k}_X = \cap_{x\in X} B^{d_k}(X - x, d_k(X))$ otherwise. The superscript d_k may be omitted if there is no risk of confusion.

Note that the notion of k-balls appears in [6] with a different meaning. Indeed, in [6], k-balls are relative to 2-way dissimilarities, and the k-ball generated by a k-element subset $A \subseteq E$, relatively to a 2-way dissimilarity d_2, is defined by

$$\mathbf{B}^{d_2}_A = \{x \in E : d_2(a, x) \leqslant \mathrm{diam}_{d_2}(A) \text{ for all } a \in A\}.$$

In fact, the k-ball $\mathbf{B}^{d_2}_A$, as defined in [6], relative to a 2-way dissimilarity d_2 is exactly the (d_k, k)-ball $B^{d_k}_A$, as defined in [10] and in the present paper, where d_k is the k-way max-extension of d_2.

Before defining quasi-ultrametrics, let us recall a well-known particular case of them, namely ultrametrics. A (2-way) dissimilarity d_2 is said to be *ultrametric* if for all x, y, z:

$$d_2(x, y) \leqslant \max\{d_2(x, z), d_2(y, z)\}.$$

Next are some characterizations of ultrametric 2-way dissimilarities, which may help in understanding the definition of quasi-ultrametrics given below.

Proposition 6 (*Diatta and Fichet [12]*). *For a 2-way dissimilarity d_2 on E, the following assertions are equivalent*:

(i) *d_2 is ultrametric*;
(ii) *for all x, y, z: the greatest two values among $d_2(x, y)$, $d_2(x, z)$ and $d_2(y, z)$ are equal*;
(iii) *for all x, y*: $\mathrm{diam}_{d_2}(B(x, d_2(x, y))) = d_2(x, y)$;
(iv) *for all x, y, u, v: $u, v \in B(x, d_2(x, y))$ implies $B(u, d_2(u, v)) \subseteq B(x, d_2(x, y))$.*

Replacing balls by k-balls in Conditions (iii) and (iv) of Proposition 6 above leads to the so-called diameter condition and inclusion condition which, together, characterize what we call quasi-ultrametric (multiway) dissimilarities [12,10]. Inclusion and diameter conditions were introduced, for 2-way dissimilarities, in [9] where inclusion condition was called "five point condition".

Definition 7. A k-way dissimilarity d_k on E is said to

(i) satisfy the *inclusion condition* if for all $X, Y \in E^*_{\leqslant k}$, $Y \subseteq B^{d_k}_X$ implies $B^{d_k}_Y \subseteq B^{d_k}_X$;
(ii) satisfy the *diameter condition* if for all $X \in E^*_{\leqslant k}$, $\mathrm{diam}_{d_k}(B^{d_k}_X) = d_k(X)$;
(iii) be *quasi-ultrametric* if it satisfies both of the inclusion and the diameter conditions.

Example 8. Fig. 2 presents three dissimilarities d_1, d'_1 and d''_1 on the set $\{i, j, k, l\}$. It is easily checked that d_1 satisfies the diameter condition; but d_1 does not satisfy the inclusion condition because $j, k \in B^{d_1}_{jl}$ whereas $i \in B^{d_1}_{jk}$ and $i \notin B^{d_1}_{jl}$.

	i	j	k	l
i	0			
j	1	0		
k	1	1	0	
l	3	2	1	0

d_1

	i	j	k	l
i	0			
j	3	0		
k	1	1	0	
l	1	1	2	0

d_1'

	i	j	k	l
i	0			
j	0	0		
k	1	1	0	
l	1	1	1	0

d_1''

Fig. 2. Three pairwise dissimilarities on the set $\{i, j, k, l\}$: d_1 satisfies the diameter but not the inclusion condition; d_1' satisfies the inclusion but not the diameter condition; d_1'' is quasi-ultrametric.

It is also easily checked that d_1' satisfies the inclusion; but d_1' does not satisfy the diameter condition because $i, j \in B_{kl}^{d_1'}$ so that $\mathrm{diam}_{d_1'}(B_{kl}^{d_1'}) > d_1'(k, l)$. The dissimilarity d_1'' is clearly quasi-ultrametric since $B_i^{d_1''} = B_j^{d_1''} = B_{ij}^{d_1''} = \{i, j\}$, for $x \neq i, j$, $B_x^{d_1''} = \{x\}$, and for $\{x, y\} \neq \{i, j\}$, $B_{xy}^{d_1''} = \{i, j, k, l\}$.

Example 9. The reader may check that the 3-way dissimilarity dis_3 defined in Example 2 is quasi-ultrametric. This can also be derived from Theorem 18 below (see Remark 19).

Let us also mention that, for 2-way dissimilarities, the diameter and inclusion conditions have been also generalized by Bertrand and Janowitz [6] into so-called k-inclusion and k-diameter conditions. According to [6], a 2-way dissimilarity d_2 is said to satisfy the k-diameter condition if $\mathrm{diam}_{d_2}(\mathbf{B}_A) = \mathrm{diam}_{d_2}(A)$ for all k-element subset A. It is then clear that the k-diameter condition, as defined in [6], for a 2-way dissimilarity, corresponds to the diameter condition, as defined in the present paper, for its k-way max-extension. On the other hand, according to [6], a 2-way dissimilarity d_2 is said to satisfy the k-inclusion condition if for all k-element subsets A_1, A_2: $A_1 \subseteq \mathbf{B}_{A_2}$ implies $\mathbf{B}_{A_1} \subseteq \mathbf{B}_{A_2}$. It is then easily observed that the k-inclusion condition, as defined in Bertrand and Janowitz [6], holds for a 2-way dissimilarity whenever its k-way max-extension satisfies the inclusion condition, as defined in the present paper. The converse does not hold because the k-inclusion condition, in the sense of [6], does not ensure that $A_1 \subseteq \mathbf{B}_{A_2}$ implies $\mathbf{B}_{A_1} \subseteq \mathbf{B}_{A_2}$ for all p-subset A_1 of E and all q-subset A_2 of E, with $1 \leqslant p, q \leqslant k$.

Given a k-way dissimilarity d_k, let us now consider the following implication introduced in [2]: for all $X \in E^*_{\leqslant k}$, $Y \in E^*_{\leqslant k-1}$ and $z \in E$,

$$z \in B_X^{d_k} \text{ implies } \mathrm{diam}_{d_k}(X \cup Y + z) \leqslant \mathrm{diam}_{d_k}(X \cup Y).$$

This implication, which we will refer to as the $2k$*-point implication*, is, in fact, a generalization of a four-point implication introduced earlier by Bandelt [1]

$$\max\{d_2(x_1, z), d_2(x_2, z)\} \leqslant d_2(x_1, x_2) \text{ implies}$$
$$d_2(y, z) \leqslant \max\{d_2(x_1, y), d_2(x_2, y), d_2(x_1, x_2)\},$$

where d_2 is a pair-wise dissimilarity.

It has been proved in [12] that quasi-ultrametric pairwise dissimilarities coincide with those which satisfy the Bandelt's four-point implication. The next proposition generalizes this result to multiway dissimilarities.

Proposition 10. *For any k-way dissimilarity d on E, the following two conditions are equivalent.*

(i) *d is quasi-ultrametric.*
(ii) *d satisfies the $2k$-point implication.*

Proof. Assume that d satisfies the $2k$-point implication and let $X, Y \in E^*_{\leqslant k}$ with $Y \subseteq B_X^d$. Then, two cases can be distinguished: either $Y = \{y\}$ or $|Y| > 1$. Assume that $Y = \{y\}$. Then, again, two cases can be distinguished: either

$X=\{x\}$ or $|X|>1$. If $X=\{x\}$, then X, X and y satisfy the conditions of the $2k$-point implication, so that

$$\mathrm{diam}_d(X\cup Y)=\mathrm{diam}_d(X\cup X+y)\leqslant \mathrm{diam}_d(X\cup X)=d(X).$$

If $|X|>1$, then for any $Z\in X_{\leqslant k-1}$, X, Z and y satisfy the condition of the $2k$-point implication, so that

$$\mathrm{diam}_d(X\cup Y)=\mathrm{diam}_d(X\cup Z+y)\leqslant \mathrm{diam}_d(X\cup Z)=d(X).$$

Assume now that $|Y|>1$. Let $y_0\in Y$ and let $Z_0=Y-y_0$. Then X, Z_0 and y_0 satisfy the condition of the $2k$-point implication, so that

$$\mathrm{diam}_d(X\cup Y)=\mathrm{diam}_d(X\cup Z_0+y_0)\leqslant \mathrm{diam}_d(X\cup Z_0).$$

if $|Z_0|=1$, then we have seen that $\mathrm{diam}_d(X\cup Z_0)\leqslant d(X)$, so that $\mathrm{diam}_d(X\cup Y)\leqslant d(X)$. If $|Z_0|>1$, pick $y_1\in Z_0$ and let $Z_1=Z_0-y_1$. Then X, Z_1 and y_1 satisfy the condition of the $2k$-point implication, so that

$$\mathrm{diam}_d(X\cup Z_0)=\mathrm{diam}_d(X\cup Z_1+y_1)\leqslant \mathrm{diam}_d(X\cup Z_1).$$

Repeating this argument, we finally get $\mathrm{diam}_d(X\cup Y)\leqslant \mathrm{d}(X)$, which proves that d satisfies the diameter condition. To prove the inclusion condition, let $u\in B^d_Y$. Then, for any $Z\in X^*_{\leqslant k-1}$, Y, Z and u satisfy the condition of the $2k$-point implication, so that $\mathrm{diam}_d(Y\cup Z+u)\leqslant \mathrm{diam}_d(Y\cup Z)$. On the other hand, $\mathrm{diam}_d(Y\cup Z)\leqslant d(X)$ since $Y\cup Z\subseteq B^d_X$ and d satisfies the diameter condition. Hence, for any $Z\in X^*_{\leqslant k-1}$,

$$d(Z+u)\leqslant \mathrm{diam}_d(Y\cup Z+u)\leqslant \mathrm{diam}_d(Y\cup Z)\leqslant d(X).$$

Therefore, $u\in B^d_X$, proving the inclusion condition.

Conversely, assume that d is quasi-ultrametric and let $X\in E^*_{\leqslant k}$, $Y\in E^*_{\leqslant k-1}$, $z\in E$ with $z\in B^d_X$. If $\mathrm{diam}_d(X\cup Y)=d(X)$, then $Y\subseteq B^d_X$ so that, by the diameter condition,

$$\mathrm{diam}_d(X\cup Y+z)\leqslant d(X)\leqslant \mathrm{diam}_d(X\cup Y).$$

As $Y\in E^*_{\leqslant k-1}$, assume, w.l.g., that $\mathrm{diam}_d(X\cup Y)=d(Z\cup T)$, where $Z\in X^*_{\leqslant p}$ and $T\in Y^*_{\leqslant k-p}$. Then $\mathrm{diam}_d(Z\cup T\cup Y)\leqslant d(Z\cup T)$ and, likewise, $\mathrm{diam}_d(Z\cup T\cup X)\leqslant d(Z\cup T)$. Thus $Y\subseteq B^d_{Z\cup T}$ and $X\subseteq B^d_{Z\cup T}$. Then, by the inclusion condition, $z\in B^d_{Z\cup T}$ since $z\in B^d_X$. Therefore, by the diameter condition,

$$\mathrm{diam}_d(X\cup Y+z)\leqslant d(Z\cup T)\leqslant \mathrm{diam}_d(X\cup Y),$$

as required. □

3. Description-meet compatibility

In this section, we place ourselves in a so-called *meet-closed description context*. That is a context consisting of a finite nonempty entity set E whose elements are described in a complete meet-semilattice $\mathcal{D}$, by means of a descriptor δ. We will denote such a context as a triple $\mathbb{K}=(E,\mathcal{D},\delta)$ where E stands for the entity set, $\mathcal{D}:=(D,\leqslant)$ the entity description space, and δ the descriptor that associates to each entity $x\in E$ its description $\delta(x)$ in $\mathcal{D}$.

In all what follows, E will denote a finite nonempty entity set, $\mathcal{D}$ a complete meet-semilattice, δ a descriptor that maps E into $\mathcal{D}$, and $\mathbb{K}$ the meet-closed description context $(E,\mathcal{D},\delta)$.

Example 11. Consider Table 2 presenting five visitors of a given Web site, described by three attributes: LiLo, NoLi, ReSu, where LiLo(x) is the login–logout time interval of visitor x within the interval [0, 24], NoLi(x) is the number of times visitor x logs in at LiLo(x) interval during a given fixed period, and ReSu(x) is the subjects requested by x during a session; requested subjects are sets of subjects from: Arts & Humanities (AH), Business & Economy (BE), Computers & Internet (CI), News & Media (NM), Recreation & Sports (RS), Science & Health (SH), Society & Culture (SC).

Then Table 2 can be seen as representing a meet-closed description context $\mathbb{K}_2:=(E_2,\mathcal{D}_2,\delta_2)$ where E_2 is the set $\{1,2,3,4,5\}$, $\mathcal{D}_2$ the direct product of three partially ordered sets (posets): the set (FUCI([0, 24]), $\subseteq$) of finite unions

Table 2
Example meet-closed description context

	LiLo	NoLi	ReSu
1	0–2	30	CI,RS
2	21–24	35	AH,NM,SC
3	0–3	40	AH,BE,CI,RS
4	22–24	35	AH,SC
5	12–14	30	BE,NM

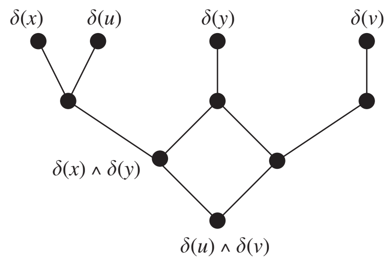

Fig. 3. A part of entity description space.

of closed intervals of [0, 24] endowed with the set inclusion order, the set $(|[30; 40]|, \leqslant)$ of integers from 30 to 40, endowed with the integer usual order, and the powerset $(\mathcal{P}(S), \subseteq)$ of the set $S = \{AH, BE, CI, NM, RS, SC\}$, endowed with the set inclusion order, and $\delta_2(x) = (\mathrm{LiLo}(x), \mathrm{NoLi}(x), \mathrm{ReSu}(x))$.

The description-meet compatibility defined below has been introduced in [13] in the case of pairwise dissimilarities. It uses the notion of valuation on a poset.

A *valuation* on a poset $(P, \leqslant)$ is a map $h : P \to \mathbb{R}_+$ such that $h(x) \leqslant h(y)$ when $x \leqslant y$. A *strict* valuation will then be a valuation h such that $x < y$ implies $h(x) < h(y)$.

Before defining the description-meet compatibility, let us introduce further notation. For any $X \subseteq E$, $\delta(X)$ will denote the set of descriptions of entities belonging to X. For any positive integer k, let $\mathcal{I}_k(E)$ denote the set of meets of descriptions of nonempty subsets of E with at most k elements, i.e.,

$$\mathcal{I}_k(E) = \{\delta(x_1) \wedge \cdots \wedge \delta(x_k) : x_1, \ldots, x_k \in E\}.$$

It may be noted that $\mathcal{I}_1(E) = \delta(E)$ and $\mathcal{I}(E) := \mathcal{I}_{|E|}(E)$ is the meet-semilattice generated by $\delta(E)$. A k-way dissimilarity d on E ($k \geqslant 2$) will be said to be δ-*meet compatible* if there exists a valuation h on $\mathcal{I}_k(E)$ with which it is δ-meet compatible, i.e., such that

$$d(X) \leqslant d(Y) \iff h(\inf\ \delta(X)) \geqslant h(\inf\ \delta(Y)),$$

for $X, Y \in E^*_{\leqslant k}$. If h is a strict valuation, d will be said to be *strictly* δ-meet compatible.

Remark 12. The reader may observe that when $\mathcal{D}$ is a complete join-semilattice, a dual compatibility condition, say δ-join compatibility, can be defined by reversing the right-hand side inequality in the above equivalence and replacing meets by joins.

Description-meet compatibility is a kind of natural agreement expressing the following fact: the lower the meet of descriptions of entities in X, the larger the dissimilarity degree of X.

To fix the ideas, assume that a part of entity description space is that depicted in Fig. 3. Then any δ-meet compatible pairwise dissimilarity d must satisfy the following inequalities: $d(x, u) \leqslant d(y, u) = d(x, y) \leqslant d(u, v) = d(x, v)$, $d(y, v) \leqslant d(u, v), \ldots$

Remark 13. If d is a strictly δ-meet compatible (multiway) dissimilarity, then $\delta(x) < \delta(y)$ implies $d(y) < d(x)$. This is why we drop out the condition $d(x) = 0$, since it is very likely to happen that two entities x and y satisfy $\delta(x) < \delta(y)$.

Example 14. Consider the meet-closed description context $\mathbb{K}_2$ defined in Example 11. Define a multiway dissimilarity on E_2 by

$$\mathrm{dis}'(X) = 47 - \left(\lambda\left(\bigcap_{x\in X} \mathrm{LiLo}(x)\right) + \min_{x\in X} \mathrm{NoLi}(x) + \left|\bigcap_{x\in X} \mathrm{ReSu}(x)\right|\right),$$

where $\lambda([\alpha, \beta]) = \beta - \alpha$. For instance, $\mathrm{dis}'(1, 2, 3) = 47 - (\lambda([0, 2] \cap [21, 24] \cap [0, 3]) + \min\{30, 35, 40\} + |\{CI, RS\} \cap \{AH, NM, SC\} \cap \{AH, BE, CI, RS\}|) = 47 - (\lambda(\emptyset) + 30 + |\emptyset|) = 47 - (0 + 30 + 0) = 17$. Then dis' is strictly δ_2-meet compatible. Indeed, λ, $x \mapsto x$ and $Y \mapsto |Y|$ are strict valuations on $(\mathrm{FUCI}([0, 24]), \subseteq)$, $(|[30; 40]|, \leqslant)$ and $(\mathcal{P}, \subseteq)$, respectively. Thus h_2 defined by

$$h_2(u, v, w) = \lambda(u) + v + |w|$$

is a strict valuation on $\mathcal{D}_2$, and the fact that dis' is δ_2-meet compatible with h_2 follows from the fact that $\mathrm{dis}'(X)$ is decreasing w.r.t. $h_2(\inf\ \delta_2(X))$.

Two valuations h and h' will be said to be *equivalent* if $h(x) \leqslant h(y)$ if and only if $h'(x) \leqslant h'(y)$. Similarly, two multiway dissimilarities d and d' will be said to be *equivalent* if $d(X) \leqslant d(Y)$ if and only if $d'(X) \leqslant d'(Y)$.

It may be noted that when a multiway dissimilarity is δ-meet compatible with a given valuation h it is also δ-meet compatible with every valuation equivalent to h. Likewise, when a valuation is δ-meet compatible with a given multiway dissimilarity d it is also δ-meet compatible with every multiway dissimilarity equivalent to d. Furthermore, two multiway dissimilarities which are respectively δ-meet compatible with two equivalent valuations are equally equivalent. A similar result holds for valuations. Moreover, we have the following characterization of the equivalence class of multiway dissimilarities δ-meet compatible with a given valuation h.

Proposition 15. *A k-way dissimilarity d is δ-meet compatible with a given valuation h if and only if there is a positive real number M such that d is equivalent to the k-way dissimilarity $d_k^{h,M}$ defined by $d_k^{h,M}(X) = M - h(\inf\ \delta(X))$.*

Proof. Let M be a positive real such that $h(\delta(x)) \leqslant M$ for any $x \in E$. By definition, d is δ-meet compatible with h if and only if for all $X, Y \in E^*_{\leqslant k}$:

$$d(X) \leqslant d(Y) \iff h(\inf\ \delta(X)) \geqslant h(\inf\ \delta(Y)).$$

Now

$$h(\inf\ \delta(X)) \geqslant h(\inf\ \delta(Y)) \iff d_k^{h,M}(X) \leqslant d_k^{h,M}(Y).$$

It then becomes clear that d and h are δ-meet compatible if and only if d and $d_k^{h,M}$ are equivalent. $\square$

The next result is straightforward but instrumental.

Proposition 16. *Let d be a δ-meet compatible k-way dissimilarity and let h be a valuation δ-meet compatible with d. Then*

(i) $d(X) = d(Y)$ *if and only if* $h(\inf\ \delta(X)) = h(\inf\ \delta(Y))$.
(ii) $\inf\ \delta(X) = \inf\ \delta(Y)$ *implies* $d(X) = d(Y)$.

Proof. Assertion (i) straightly derives from δ-meet compatibility of d and h. Assertion (ii) is then a consequence of the fact that $\inf\ \delta(X) = \inf\ \delta(Y)$ implies $h(\inf\ \delta(X)) = h(\inf\ \delta(Y))$. $\square$

Before outlining the relationship between quasi-ultrametricity and description-meet compatibility, let us recall the following technical notion: the *breadth* of a meet-semilattice $(P, \leqslant)$ is the least positive integer k such that the meet of any $(k+1)$ elements of P is always the meet of k elements among these $k+1$ [7]. Having noticed this, we agree to say that a subset Q of a meet-semilattice is of breadth k if k is the least positive integer such that for any $(k+1)$-element subset W of Q there is $w \in W$ such that $\inf(W-w) \leqslant w$.

Example 17. Consider the dataset given in Table 1 as presenting a meet-closed description context $\mathbb{K}_1 := (E_1, \mathcal{D}_1, \delta_1)$, where E_1 is the set of five items and $\mathcal{D}_1$ the boolean lattice $\{0,1\}^7$; for instance $\delta_1(\text{brd})=(1,0,1,0,1,1,1)$. Then $\delta_1(E_1)$ is of breadth at least 3 since $\inf\ \delta_1(\{\text{brd, chs, mlk}\}) = (0,0,0,0,0,0,1)$, which is different from either of $\delta_1(\text{brd}) \wedge \delta_1(\text{chs})=(1,0,1,0,0,0,1)$, $\delta_1(\text{brd})\wedge\delta_1(\text{mlk})=(0,0,0,0,1,0,1)$ and $\delta_1(\text{chs})\wedge\delta_1(\text{mlk})=(0,1,0,0,0,0,1)$. Moreover, $\inf\ \delta_1(\{\text{brd, btr, chs, egg}\}) = \inf \delta_1(\{\text{brd, btr, chs, mlk}\}) = \inf\ \delta_1(\{\text{brd, btr, chs}\})$, $\inf \delta_1(\{\text{brd, btr, egg, mlk}\}) = \inf\ \delta_1(\{\text{brd, chs, egg, mlk}\}) = \inf \delta_1(\{\text{brd, egg, mlk}\})$, and $\inf\ \delta_1(\{\text{btr, chs, egg, mlk}\}) = \inf \delta_1(\{\text{btr, chs, egg}\})$,so that $\delta_1(E_1)$ is of breadth 3.

We now go on proving the result showing the existence of an integer $k \geqslant 2$ such that any strictly δ-meet compatible k-way dissimilarity on E is quasi-ultrametric.

Theorem 18. (i) *If $\delta(E)$ is of breadth one, then every strictly δ-meet compatible 2-way dissimilarity on E is ultrametric.*
(ii) *If $\delta(E)$ is of breadth $k \geqslant 2$, then every strictly δ-meet compatible k-way dissimilarity on E is quasi-ultrametric.*

Proof. (i) Let d_2 be a strictly δ-meet compatible 2-way dissimilarity on E and let $x, y, z \in E$. As $\delta(E)$ is of breadth one, any two elements $\delta(u), \delta(v) \in \delta(E)$ are such that either $\delta(u) \leqslant \delta(v)$ or $\delta(v) \leqslant \delta(u)$. Assume, w.l.g., that $\delta(x) \leqslant \delta(y) \leqslant \delta(z)$. Then

$$\delta(x) \wedge \delta(y) = \delta(x) \wedge \delta(z) = \delta(x)$$

so that, according to Proposition 16 (ii), $d_2(x, y) = d_2(x, z)$. On the other hand,

$$\delta(x) = \delta(x) \wedge \delta(y) \leqslant \delta(y) = \delta(y) \wedge \delta(z)$$

so that, by δ-meet compatibility of d_2, $d_2(y, z) \leqslant d_2(x, y)$. Then the greatest two values among $d_2(x, y)$, $d_2(x, z)$ and $d_2(y, z)$ are equal, and the result follows from Proposition 6.

(ii) Let d_k be a strictly δ-meet compatible k-way dissimilarity on E. According to Proposition 10, d_k is quasi-ultrametric if and only if it satisfies the $2k$-point implication. To show that d_k satisfies the $2k$-point implication, let $X \in E^*_{\leqslant k}$, $Y \in E^*_{\leqslant k-1}$, and $z \in B^{d_k}_X$. Then we need to prove that $\mathrm{diam}_{d_k}(X \cup Y + z) \leqslant \mathrm{diam}_{d_k}(X \cup Y)$. As $\delta(E)$ is of breadth k, we claim that $\inf\ \delta(X) \leqslant \delta(z)$. Indeed, either $(X+z) \in E^*_{\leqslant k}$ or not. Assume that $(X+z) \in E^*_{\leqslant k}$. Then, as $z \in B^{d_k}_X$, we have $d_k(X+z) \leqslant d_k(X)$ so that, by δ-meet compatibility, $h(\inf\ \delta(X)) \leqslant h(\inf\ \delta(X+z))$ for any strict valuation h δ-meet compatible with d_k. On the other hand, $h(\inf \delta(X+z)) \leqslant h(\inf \delta(X))$ since $\delta(X+z) \leqslant \delta(X)$. Then $h(\inf\ \delta(X+z)) = h(\inf\ \delta(X))$, so that $\inf\ \delta(X+z) = \inf\ \delta(X)$, since h is a strict valuation. Therefore,

$$\inf\ \delta(X) = \inf\ \delta(X+z) = \inf\ \delta(X) \wedge \delta(z) \leqslant \delta(z).$$

Assume now that $(X+z) \notin E^*_{\leqslant k}$. Then $|X+z| = k+1$ since $X \in E^*_{\leqslant k}$. Therefore, as $\delta(E)$ is of breadth k, we have necessarily: (a) $\inf\ \delta(X) \leqslant \delta(z)$, or (b) there exists $x \in X$ such that $\inf\ \delta((X-x)+z) \leqslant \delta(x)$. It follows from (b) that

$$\begin{aligned}\inf\ \delta((X-x)+z) &= \inf\ \delta(X-x) \wedge \inf\ \delta((X-x)+z)\\ &\leqslant \inf\ \delta(X-x) \wedge \inf\ \delta(x) = \inf\ \delta(X).\end{aligned}$$

Then, by δ-meet compatibility of d_k, $d_k(X) \leqslant d_k((X-x)+z)$. Hence $d_k((X-x)+z) = d_k(X)$ since it is otherwise assumed that $z \in B^{d_k}_X$, i.e., $d_k((X-x)+z) \leqslant d_k(X)$. Thus, by Proposition 16 (i), $h(\inf\ \delta((X-x)+z) = h(\inf\ \delta(X))$ for any strict valuation h δ-meet compatible with d_k. Hence $\inf\ \delta((X-x)+z) = \inf\ \delta(X)$, from which it follows that:

$$\inf\ \delta(X) = \inf\ \delta((X-x)+z) = \inf\ \delta(X-x) \wedge \delta(z) \leqslant \delta(z).$$

Therefore, in all cases, $\inf\ \delta(X) \leqslant \delta(z)$, as claimed. Then for any $Z \in (X \cup Y)^*_{\leqslant k-1}$, we have $\inf\ \delta(X \cup Z) \leqslant \inf\ \delta(Z+z)$. Now, as $\delta(E)$ is of breadth k, for any $Z \in (X \cup Y)^*_{\leqslant k-1}$, there is $T \in (X \cup Z)^*_{\leqslant k-1}$ such that $\inf\ \delta(T) = \inf\ \delta(X \cup Z)$, so that $\mathrm{diam}_{d_k}(X \cup Z) = d_k(T)$. Hence, as $\inf\ \delta(T) = \inf\ \delta(X \cup Z) \leqslant \inf\ \delta(Z + z)$, we have, by δ-meet compatibility of d_k, $d_k(Z + z) \leqslant d_k(T)$. Then for any $Z \in (X \cup Y)^*_{\leqslant k-1}$

$$d_k(Z + z) \leqslant \mathrm{diam}_{d_k}(X \cup Z) \leqslant \mathrm{diam}_{d_k}(X \cup Y).$$

Therefore,

$$\mathrm{diam}_{d_k}(X \cup Y + z) \leqslant \mathrm{diam}_{d_k}(X \cup Y),$$

proving that d_k satisfies the $2k$-point implication. □

The converse of Theorem 18 does clearly not hold since, for $k \geqslant 2$, every constant (one-valued) k-way dissimilarity on E is quasi-ultrametric but never strictly δ-meet compatible, regardless to the descriptor δ. Indeed, otherwise, we would have, for all $x, y \in E$, $\delta(x) = \delta(y)$ so that $\delta(E)$ would be a singleton, hence of breadth one.

Remark 19. As claimed in Example 9, it follows from Theorem 18 that the 3-way dissimilarity dis_3 defined in Example 2 is quasi-ultrametric. Indeed, on the one hand, as observed in Example 17, $\delta_1(E_1)$ is of breadth 3. On the other hand, for each k such that $2 \leqslant k \leqslant 5$, dis_k is strictly δ_1-meet compatible with the valuation h_1 defined on $\mathcal{D}_1$ by letting $h_1(x)$ be the number of ones occurring in x.

The entity set E being finite, there is an integer $k \geqslant 1$ such that k is the breadth of $\delta(E)$. Moreover, as any pairwise ultrametric dissimilarity is quasi-ultrametric, we derive the following from Theorem 18.

Corollary 20. *There is an integer $k \geqslant 2$ such that any strictly* δ-meet compatible *k-way dissimilarity on E is quasi-ultrametric*.

Following [10], a k-way dissimilarity d will be said to be ultrametric if for all $X \in E^*_{\leqslant k}$ and $x \in E$:

$$d(X) \leqslant \max_{Y \in X^*_{\leqslant k-1}} d(Y + x).$$

It should be noticed that any canonical restriction of a strictly δ-meet compatible multiway dissimilarity is also strictly δ-meet compatible. Then, when $\delta(E)$ is of breadth one, for an integer $p \geqslant 2$, the 2-way canonical restriction d_2 of any strictly δ-meet compatible p-way dissimilarity d_p is ultrametric. Now, by Proposition 33 (Section 4), d_p is equivalent to the p-way max-extension of d_2, which, by Corollary 26 (Section 4), is a p-way ultrametric dissimilarity. Therefore, d_p is ultrametric since, for an integer $q \geqslant 2$, any q-way dissimilarity equivalent to a q-way ultrametric one is also ultrametric. Finally, when $\delta(E)$ is of breadth one, Theorem 18 (i) extends to ultrametric multiway dissimilarities:

Theorem 21. *If $\delta(E)$ is of breadth one*, *then for $k \geqslant 2$*, *every strictly* δ-meet compatible *k-way dissimilarity on E is ultrametric*.

4. Relationships between multiway dissimilarities

In this section, we outline properties that multiway dissimilarities transmit to or inherit from their max-extensions or canonical restrictions. We begin with results from the author's earlier paper.

Let p and q be two positive integers such that $2 \leqslant p \leqslant q$. The following proposition states that k-balls relative to a q-way dissimilarity are, respectively, contained in those with the same generators, relative to its p-way canonical restriction.

Proposition 22 (*Diatta [10]*). *Let d_q be a q-way dissimilarity on E and let d_p be its p-way canonical restriction. Then for all $X \in E^*_{\leqslant p}$, $B_X^{d_q} \subseteq B_X^{d_p}$.*

The next result states that the set of k-balls relative to a p-way dissimilarity contains that relative to its q-way max-extension.

Proposition 23 (*Diatta [10]*). *Let d_p be a p-way dissimilarity and let d_q be its q-way max-extension. Then for all $X \in E^*_{\leqslant q}$, $B_X^{d_q} = B_Y^{d_p}$, where $Y \in X^*_{\leqslant p}$ is such that $d_p(Y) = \mathrm{diam}_{d_p}(X)$.*

As a direct consequence of Proposition 23, we have the following corollary, also observed in [10].

Corollary 24. *Any q-way max-extension of a p-way quasi-ultrametric dissimilarity is quasi-ultrametric.*

The last result we recall from [10] states that every q-way ultrametric dissimilarity is the q-way max-extension of its p-way canonical restriction.

Proposition 25 (*Diatta [10]*). *Every q-way ultrametric dissimilarity is the q-way max-extension of its p-way canonical restriction which, in turn, is ultrametric.*

From Propositions 23 and 25, we can derive that for any p-way ultrametric dissimilarity d_p, and for all $X \in E^*_{\leqslant p}$: $B_X^{d_p} = B_{xy}^{d_2}$, where d_2 is the 2-way canonical restriction of d_p, and where x, y are such that $d_2(x, y) = \mathrm{diam}_{d_2}(X)$. Moreover, as observed in [10], $B_X^{d_p} = B(x, d_2(x, y))$, since, by Proposition 6 (iv), for all u, v: $B_{uv}^{d_2} = B(u, d_2(u, v))$ because d_2 is ultrametric. The following is then easily observed.

Corollary 26. *Any q-way max-extension of a p-way ultrametric dissimilarity is ultrametric.*

The next result gives further information about relationships between k-balls relative to a q-way dissimilarity and those relative to its p-way canonical restriction: it specifies entity subsets which generate the same k-balls relatively to both the p-way and the q-way dissimilarities.

Proposition 27. *Let d_q be a q-way dissimilarity and let d_p be its p-way canonical restriction. Then for all $X \in E^*_{\leqslant p-1}$, $B_X^{d_q} = B_X^{d_p}$.*

Proof. Let $X \in E^*_{\leqslant p-1}$ and let $x \in E$. Then $x \in B_X^{d_q}$ if and only if $d_q(X+x) \leqslant d_q(X)$. Similarly, $x \in B_X^{d_p}$ if and only if $d_p(X+x) \leqslant d_p(X)$. Now, $d_p(X+x) = d_q(X+x)$ and $d_p(X) = d_q(X)$, since d_p is the p-way canonical restriction of d_q. Therefore, $B_X^{d_q} = B_X^{d_p}$, as required. □

The following result shows that if both of two multiway dissimilarities are δ-meet compatible with the same valuation, then one of them if equivalent to the canonical restriction of the other.

Proposition 28. *If d_p and d_q are respectively a p-way dissimilarity and q-way dissimilarity both δ-meet compatible with the same valuation, then d_p is equivalent to the p-way canonical restriction of d_q.*

Proof. Let h be a valuation δ-meet compatible with both d_p and d_q. Then, by Proposition 15, there is a positive real number M such that d_p and d_q are equivalent to $d_p^{h,M}$ and $d_q^{h,M}$, respectively. Now $d_p^{h,M}$ is the p-way canonical restriction of $d_q^{h,M}$, proving the required equivalence between d_p and the p-way canonical restriction of d_q. □

It may be noticed that Proposition 28 remains valid if "*the same valuation*" is replaced by "*two equivalent valuations*".

Remark 29. Two equivalent k-way dissimilarities d and d' have the same k-balls; i.e., for all $X \in E^*_{\leqslant k}$, $B_X^d = B_X^{d'}$. Moreover, for all $Y \subseteq E$, $\mathrm{diam}_d(Y) = d(Z)$ if and only if $\mathrm{diam}'_d(Y) = d'(Z)$.

The next result shows that when $\delta(E)$ is of breadth p, then a k-ball relative to a p-way δ-meet compatible dissimilarity, generated by a subset X, is relative to a $(|X|+1)$-way δ-meet compatible dissimilarity if $|X| < p$, or to a $(p+1)$-way δ-meet compatible one otherwise.

Proposition 30. *Let h be a valuation on $\mathcal{I}(E)$ and, for any integer $p \geqslant 2$, let d_p be a p-way dissimilarity on E, δ-meet compatibel with h. If $\delta(E)$ is of breadth $k \geqslant 2$, then for all $X \in E^*_{\leqslant k-1}$: $B_X^{d_{|X|+1}} = B_X^{d_k}$. Moreover, for all $q \geqslant k$ and all $X \in E^*_{\leqslant k}$: $B_X^{d_q} = B_X^{d_k}$.*

Proof. To prove the first assertion, let $X \in E^*_{\leqslant k-1}$. If $|X| = k-1$, then $d_{|X|+1} = d_k$, so that $B_X^{d_{|X|+1}} = B_X^{d_k}$. Assume that $|X| < k-1$ and let $d_{k,|X|+1}$ denote the $(|X|+1)$-way canonical restriction of d_k. Then, according to Proposition 28, $d_{|X|+1}$ is equivalent to $d_{k,|X|+1}$ since both $d_{|X|+1}$ and d_k are δ-meet compatible with h. Now, by Proposition 27, $B_X^{d_k} = B_X^{d_{k,|X|+1}}$. Therefore, by Remark 29, $B_X^{d_k} = B_X^{d_{|X|+1}}$. To prove the second assertion, let $X \in E^*_{\leqslant k}$ and let $d_{q,k}$ denote the k-way canonical restriction of d_q. Observe first that, by Proposition 28, d_k is equivalent to $d_{q,k}$. If $|X| < k$, then, by Proposition 27, $B_X^{d_q} = B_X^{d_{q,k}}$. On the other hand, by Remark 29, $B_X^{d_k} = B_X^{d_{q,k}}$, proving that $B_X^{d_q} = B_X^{d_k}$. Assume that $|X| = k$. By Proposition 22, $B_X^{d_q} \subseteq B_X^{d_{q,k}}$. Moreover, by Remark 29, $B_X^{d_{q,k}} = B_X^{d_k}$, so that $B_X^{d_q} \subseteq B_X^{d_k}$. The converse inclusion derives from the fact that $\delta(E)$ is of breadth k. Indeed, let $y \in B_X^{d_k}$, i.e., $y \in B_X^{d_{q,k}}$ or, in other words,

$$d_q(Y+y) = d_{q,k}(Y+y) \leqslant d_{q,k}(X) = d_q(X)$$

for any $Y \in X^*_{\leqslant k-1}$. As $\delta(E)$ is assumed to be of breadth k, there is $Z \in (X+y)^*_{\leqslant k}$ such that $\inf\ \delta(X+y) = \inf\ \delta(Z)$, so that, by proposition 16 (ii), $d_q(X+y) = d_q(Z)$. Two cases can then be distinguished: either $Z = X$ or $Z = Y+y$ for some $Y \in X^*_{\leqslant k-1}$. On the one hand, if $Z = X$, then $d_q(X+y) = d_q(X)$, so that $y \in B_X^{d_q}$. On the other hand, if $Z = Y+y$ for some $Y \in X^*_{\leqslant k-1}$, then

$$d_q(X+y) = d_q(Y+y) = d_{q,k}(Y+y) \leqslant d_{q,k}(X) = d_q(X),$$

so that $y \in B_X^{d_q}$. Therefore $B_X^{d_k} \subseteq B_X^{d_q}$, proving the required equality. □

As a consequence of Proposition 30, when $\delta(E)$ is of breadth p, each k-ball relative to a p-way δ-meet compatible dissimilarity can be computed in $0(|E|)$. The next result shows that the k-ball generated by a subset X relatively to a $(|X|+1)$-way δ-meet compatible dissimilarity is equal to the k-ball generated by some proper subset of X relatively to a $|X|$-way δ-meet compatible dissimilarity.

Proposition 31. *Let h be a strict valuation on $\mathcal{I}(E)$ and, for any $p \geqslant 2$, let d_p be a p-way dissimilarity δ-meet compatible with h. Then for all $X \in E^*_{\leqslant p+1}$, $B_X^{d_{p+2}} = B_Y^{d_{p+1}}$ for any $Y \in X^*_{\leqslant p}$ such that $d_{p+1}(X) = d_p(Y)$.*

Proof. Let $X \in E^*_{\leqslant p+1}$ and let $Y \in X^*_{\leqslant p}$ such that $d_{p+1}(X) = d_p(Y)$. Observe first that, according to Proposition 15, there is a positive real number M such that d_p is equivalent to $d_p^{h,M}$ for each $p \geqslant 2$. Observe also that $y \in B_X^{d_{p+2}^{h,M}}$ if and only if $d_{p+2}^{h,M}(X+y) \leqslant d_{p+2}^{h,M}(X) = d_{p+1}^{h,M}(X)$. Likewise, $y \in B_Y^{d_{p+1}^{h,M}}$ if and only if $d_{p+1}^{h,M}(Y+y) \leqslant d_{p+1}^{h,M}(Y) = d_p^{h,M}(Y)$. Moreover, $d_{p+1}(X) = d_p(Y)$ implies $d_{p+1}^{h,M}(X) = d_p^{h,M}(Y)$. Then, by Proposition 16 (ii), $h(\inf\ \delta(X)) = h(\inf\ \delta(Y))$. Thus $\inf\ \delta(X) = \inf\ \delta(Y)$ since, on the one hand, $\inf\ \delta(X) \leqslant \inf\ \delta(Y)$, and, on the other hand, h is a strict valuation. Then $\inf\ \delta(X+y) = \inf\ \delta(Y+y)$, so that, again by Proposition 16 (ii), $d_{p+2}^{h,M}(X+y) = d_{p+1}^{h,M}(Y+y)$. Therefore, $y \in B_X^{d_{p+2}^{h,M}}$ if and only if $y \in B_Y^{d_{p+1}^{h,M}}$, i.e., $B_X^{d_{p+2}^{h,M}} = B_Y^{d_{p+1}^{h,M}}$. Now, by Proposition 29, $B_X^{d_{p+2}} = B_X^{d_{p+2}^{h,M}}$ and $B_Y^{d_{p+1}} = B_Y^{d_{p+1}^{h,M}}$, proving the required equality. □

The following proposition shows that max-extensions of two equivalent multiway dissimilarities are equivalent too.

Proposition 32. *If d_k and d'_k are two equivalent k-way dissimilarities, then for all $l \geqslant k$, the l-way max-extensions of d_k and d'_k are equivalent too.*

Proof. Let d_k and d'_k be two equivalent k-way dissimilarities, i.e., such that for any $X, Y \in E^*_{\leqslant k}$, $d_k(X) \leqslant d_k(Y)$ if and only if $d'_k(X) \leqslant d'_k(Y)$. Let d_l and d'_l be the l-way max-extensions of d_k and d'_k, respectively, and let $X, Y \in E^*_{\leqslant l}$. Assume that $d_l(X) \leqslant d_l(Y)$. Assume also that $d_l(X) = d_k(X')$ and $d_l(Y) = d_k(Y')$, where $X' \in X^*_{\leqslant k}$ and $Y' \in Y^*_{\leqslant k}$, with

$$d_k(X') = \max_{Z \in X^*_{\leqslant k}} d_k(Z) \quad \text{and} \quad d_k(Y') = \max_{Z \in Y^*_{\leqslant k}} d_k(Z).$$

Then, on the one hand,

$$d'_k(X') = \max_{Z \in X^*_{\leqslant k}} d'_k(Z) \quad \text{and} \quad d'_k(Y') = \max_{Z \in Y^*_{\leqslant k}} d'_k(Z).$$

because d'_k is equivalent to d_k. Hence $d'_l(X) = d'_k(X')$ and $d'_l(Y) = d'_k(Y')$. On the other hand, $d'_k(X') \leqslant d'_k(Y')$, since d_k and d'_k are equivalent and it is assumed that $d_k(X') = d_l(X) \leqslant d_l(Y) = d_k(Y')$. Therefore, $d'_l(X) \leqslant d'_l(Y)$, as required. □

The next result shows that for two integers p and q greater than or equal to the breadth of $\delta(E)$, if both of a p-way dissimilarity and a q-way one are δ-meet compatible with the same valuation, then one of them is equivalent to a max-extension of the other.

Proposition 33. *Let $\delta(E)$ be of breadth $k \geqslant 1$ and let h be a valuation on $\mathcal{I}(E)$. For any integer $l \geqslant 2$, let d_l be an l-way dissimilarity δ-meet compatible with h. Then for all positive integers p and q such that* $\max\{2, k\} \leqslant p \leqslant q$, *$d_q$ is equivalent to the q-way max-extension of d_p.*

Proof. Let h be a valuation δ-meet compatible with both d_p and d_q. Then, according to Proposition 15, there is a positive real number M such that d_p and d_q are equivalent to $d_p^{h,M}$ and $d_q^{h,M}$, respectively. Let $X \in E^*_{\leqslant q}$. Then, as $\delta(E)$ is of breadth $k \leqslant p$, there is $Y \in X^*_{\leqslant p}$ such that $\inf\ \delta(X) = \inf\ \delta(Y)$. Then for all $Z \in X^*_{\leqslant p}$, $\inf\ \delta(Y) \leqslant \inf\ \delta(Z)$. Hence $h(\inf\ \delta(Y)) = \min_{Z \in X^*_{\leqslant p}} h(\inf\ \delta(Z))$. Therefore,

$$d_p^{h,M}(Y) = M - h(\inf\ \delta(Y)) = M - \min_{Z \in X^*_{\leqslant p}} h(\inf\ \delta(Z)) = \max_{Z \in X^*_{\leqslant p}} d_p^{h,M}(Z).$$

Now,

$$d_q^{h,M}(X) = M - h(\inf\ \delta(X)) = M - h(\inf\ \delta(Y)) = d_p^{h,M}(Y).$$

Then $d_q^{h,M}$ is the q-way max-extension of $d_p^{h,M}$. Hence, by Proposition 32, d_q is equivalent to the q-way max-extension of d_p, since d_p is equivalent to $d_p^{h,M}$. □

According to Remark 29, any k-way dissimilarity equivalent to a quasi-ultrametric k-way one is also quasi-ultrametric. Then we have the following:

Proposition 34. *If $\delta(E)$ is of breadth $k \geqslant 2$, then for all $l \geqslant k$, every strictly δ-meet compatible l-way dissimilarity is quasi-ultrametric.*

Proof. Let d_l be a l-way strictly δ-meet compatible dissimilarity and let h be a valuation δ-meet compatible with d_l. Then, by Theorem 18, the k-way dissimilarity $d_k^{h,M}$, where $M = \max_{x \in E} h(\delta(x))$, is quasi-ultrametric. On the other hand, by Proposition 33, d_l is equivalent to the l-way max-extension of $d_k^{h,M}$. Therefore, d_l is quasi-ultrametric, as any max-extension of a multiway quasi-ultrametric dissimilarity is also quasi-ultrametric [10, Proposition 4]. □

5. Examples of description-meet compatible multiway dissimilarities

5.1. The canonical description-meet compatible multiway dissimilarity

The canonical description-meet compatible multiway dissimilarity is a strictly description-meet compatible multiway dissimilarity which one can derive from any meet-closed description context. Consider the context $\mathbb{K} = (E, \mathcal{D}, \delta)$ and let h_c be the map defined on $\mathcal{I}(E)$ by

$$h_c(\omega) = |\{\omega' \in \mathcal{I}(E) : \omega' \leqslant \omega\}|,$$

i.e., the number of elements of $\mathcal{I}(E)$ which are less than or equal to ω. It is then easily observed that h_c is a strict valuation on $\mathcal{D}$. The *canonical* δ-meet compatible multiway dissimilarity on E is the multiway dissimilarity $d^{h_c,M}$ defined by

$$d^{h_c,M}(X) = M - h_c(\inf\ \delta(X)),$$

where $M = \max_{x \in E} h_c(\delta(x))$.

5.2. The Ochiaï's multiway dissimilarity

The Ochiaï's [16] 2-way dissimilarity is well known for qualitative-attribute/object data. Assume entities in E be described by a set of p qualitative attributes, each of these attributes taking its values in a finite set of modalities. For an attribute a, let $\mathrm{dom}(a)$ denote the domain of a. If the modalities of $\mathrm{dom}(a)$ are listed in a given fixed linear ordering, then each of them may be represented by a boolean $|\mathrm{dom}(a)|$-vector of zeros and ones with a single one at its corresponding rank. Then an attribute a may be regarded as mapping the set E into the finite lattice $\{0, 1\}^{|\mathrm{dom}(a)|}$, so that entities description space is the lattice $\{0, 1\}^q$, where $q = \sum_{a \in A} |\mathrm{dom}(a)|$. Thus the description of an entity x is the q-vector of zeros and ones $\delta(x) = (x^i)_{i=1,\ldots,q}$ where $x^i = 1$ if and only if the modality at the ith rank is observed on x.

Let h be the map defined on $\{0, 1\}^q$ by letting $h(\omega)$ be the number of ones occurring in ω. Then h is a strict index which has the following obvious properties.

- For all $X \subseteq E$, $h(\inf\ \delta(X))$ is the number of modalities simultaneously observed on entities belonging to X.
- For all $x \in E$, $h(\delta(x)) = p$ since for each attribute, exactly one modality of this attribute is observed on x.

The Ochiaï's multiway dissimilarity is defined by

$$d^{\mathrm{O}}(X) = \left[2\left(1 - \frac{h(\inf \delta(X))}{\left[\prod_{x \in X} h(\delta(x))\right]^{1/|X|}}\right)\right]^{1/2} = \left[2\left(1 - \frac{h(\inf\ \delta(X))}{p}\right)\right]^{1/2}$$

The δ-meet compatibility of d^{O} derives from the fact that d^{O} is a decreasing function of the argument $h(\inf\ \delta(X))$.

5.3. A practical δ-meet compatible multiway dissimilarity

Our last example of δ-meet compatible multiway dissimilarity is of the type $d_k^{h,M}$, where h is a simply computable valuation. Assume δ maps E into a finite Cartesian product of meet-semilattices $\mathcal{D}_i$, $1 \leqslant i \leqslant p$, as it is the case when we are concerned with usual attribute/object data. It can then be convenient to consider a valuation h_i on each $\mathcal{D}_i$, and to define h on $\prod_{i \in \{1,\ldots,p\}} \mathcal{D}_i$ by

$$h(\omega_1, \ldots, \omega_p) = \sum_{i \in \{1,\ldots,p\}} h_i(\omega_i).$$

It should be noticed that h is strict if h_i is strict for each i. The multiway dissimilarity dis' defined in Example 14 is of this kind.

6. Conclusion

We presented a class of multiway dissimilarities and studied some of their main properties. Besides the fact that multiway dissimilarities, unlike usual 2-way ones, allow global comparison of more than two entities, those dealt with in this paper agree with entity descriptions in a very natural sense expressed by a condition called δ-meet compatibility.

We have shown that there is always an integer k such that any strictly δ-meet compatible k-way dissimilarity is quasi-ultrametric. Now quasi-ultrametric (multiway) dissimilarities are among those (multiway) dissimilarities which play an important role in mathematical clustering. Indeed, they extend ultrametrics which are well known to be in bijection with indexed hierarchies [14,5]. On the other hand, they are also in bijection with indexed (k-) quasi-hierarchies [10–12] also known as indexed closed weak hierarchies of breadth at most k [2] or k-weak hierarchical representations [6]. We are then confident that the properties outlined in this paper would help for further contributions in cluster analysis.

References

[1] H.-J. Bandelt, Four point characterization of the dissimilarity functions obtained from indexed closed weak hierarchies, mathematisches Seminar, Universität Hamburg, Germany, 1992.
[2] H.-J. Bandelt, A.W.M. Dress, An order theoretic framework for overlapping clustering, Discrete Math. 136 (1994) 21–37.
[3] A. Batbedat, Les dendrogrammes des dissimilarités symétriques n-voies: comment situer les graphes symétriques n-voies, Technical report, Université de Montpellier II, France, 1993.
[4] M. Bennani, W.J. Heiser, Triadic distance models: axiomatization and least squares representation, J. Math. Psych. 41 (1997) 189–206.
[5] J.-P. Benzécri, L'Analyse des données: la Taxinomie, Dunod, Paris, 1973.
[6] P. Bertrand, M.F. Janowitz, The k-weak hierarchical representations: an extension of the indexed closed weak hierarchies, Discrete Appl. Math. 127 (2003) 199–220.
[7] G. Birkhoff, Lattice theory, third ed., Collequium Publications, vol. XXV, American Mathematical Society, Providence, RI, 1967.
[8] J.T. Daws, The analysis of free-sorting data: beyond pairwise cooccurrences, J. Classification 13 (1996) 57–80.
[9] J. Diatta, Hiérarchies faibles et quasi-hiérarchies, mémoire de DEA, Université de Provence, France, 1992.
[10] J. Diatta, Dissimilarités multivoies et généralisations d'hypergraphes sans triangles, Math. Inf. Sci. Hum. 138 (1997) 57–73.
[11] J. Diatta, B. Fichet, From Apresjan hierarchies and Bandelt–Dress weak hierarchies to quasi-hierarchies, in: E. Diday, Y. Lechevalier, M. Schader, P. Bertrand, B. Burtschy (Eds.), New Approaches in Classification and Data Analysis, Springer, Berlin, 1994, pp. 111–118.
[12] J. Diatta, B. Fichet, Quasi-ultrametrics and their 2-ball hypergraphs, Discrete Math. 192 (1998) 87–102.
[13] J. Diatta, H. Ralambondrainy, The conceptual weak hierarchy associated with a dissimilarity measure, Math. Social Sci. 44 (2002) 301–319.
[14] S.C. Johnson, Hierarchical clustering schemes, Psychometrika 32 (1967) 241–254.
[15] S. Joly, G. Le Calvé, Three-way distances, J. Classification 12 (1995) 191–205.
[16] A. Ochiaï, Zoogeographic studies on the solenoïd fishes found in Japan and its neighbouring regions, Bull. Japan Soc. Sci. Fish. 22 (1957) 526–530.